

PREDIKSI KETEPATAN KELULUSAN MAHASISWA MENGUNAKAN ALGORITMA C5.0

Lindung Budiyanoro, Mula'ab, S.Si, M.Kom, Fika Hastarita R.,S.T, M.Eng.
Program Studi Teknik Informatika, Universitas Trunojoyo
Jl. Raya Telang PO. BOX 2 Kamal, Bangkalan - 69162
e-mail: lindung@windowslive.com, mulaab@if.trunojoyo.ac.id, hastarita.fika@gmail.com

ABSTRAK

Selama ini prediksi ketepatan kelulusan mahasiswa belum pernah dilakukan di jurusan Teknik Informatika, Universitas Trunojoyo, Madura. Prediksi ketepatan kelulusan mahasiswa perlu dilakukan untuk mengetahui ketepatan kelulusan seorang mahasiswa. Mahasiswa dinyatakan lulus tepat waktu apabila menempuh jenjang pendidikan strata 1 selama 8 semester atau 4 tahun. Salah satu teknik data mining yang dapat digunakan untuk melakukan prediksi adalah dengan menggunakan metode *decision tree*. Metode ini akan menghasilkan pohon dan *rule* keputusan yang nantinya akan dijadikan sebagai acuan dalam proses prediksi. Sistem prediksi ini menggunakan algoritma C5.0 yang merupakan salah satu metode *decision tree*. Data mahasiswa yang terdiri dari 37 atribut dan 1 kelas, dihitung menggunakan rumus *information gain* untuk menghasilkan pohon dan *rule* keputusan. Dari uji coba yang telah dilakukan, terbentuk pohon keputusan yang berbeda-beda. Perbedaan ini disebabkan oleh jumlah *data training* yang digunakan pada setiap partisi data jumlahnya berbeda. Pohon keputusan partisi A menggunakan teknik *postpruning* dengan jumlah *data training* lebih besar daripada *data testing* mempunyai nilai akurasi tertinggi sebesar 90.32 %, sehingga pohon keputusan ini dipilih sebagai acuan dalam proses prediksi.

Kata kunci : *Decision tree, rule, algoritma C5.0, information gain, postpruning.*

ABSTRACT

During the graduation prediction accuracy has never been done in the Department of Informatics Engineering, Trunojoyo University, Madura. Prediction accuracy graduation needs to be done to determine the accuracy of the graduation of a student. Students graduated on time in an undergraduate education take 1 for 8 semesters or 4 years. One of data mining techniques that can be used to make predictions is to use a decision tree method. This method will result in a decision tree and a rule that will be used as reference in the prediction process. This prediction system using the C5.0 algorithm, which is one method of decision tree. The data consisted of 37 students and one class attributes, information gain is calculated using the formula to generate decision trees and rule. Of trials that have been conducted, a decision tree formed different. This difference is caused by the amount of training data used in any number of different data partitions. A decision tree partitions using postpruning technique training with larger amounts of data than the data testing has the highest accuracy value of 90.32%, so the decision tree is chosen as a reference in the prediction process.

Keywords: *Decision tree, rule, C5.0 algorithm, information gain, postpruning.*